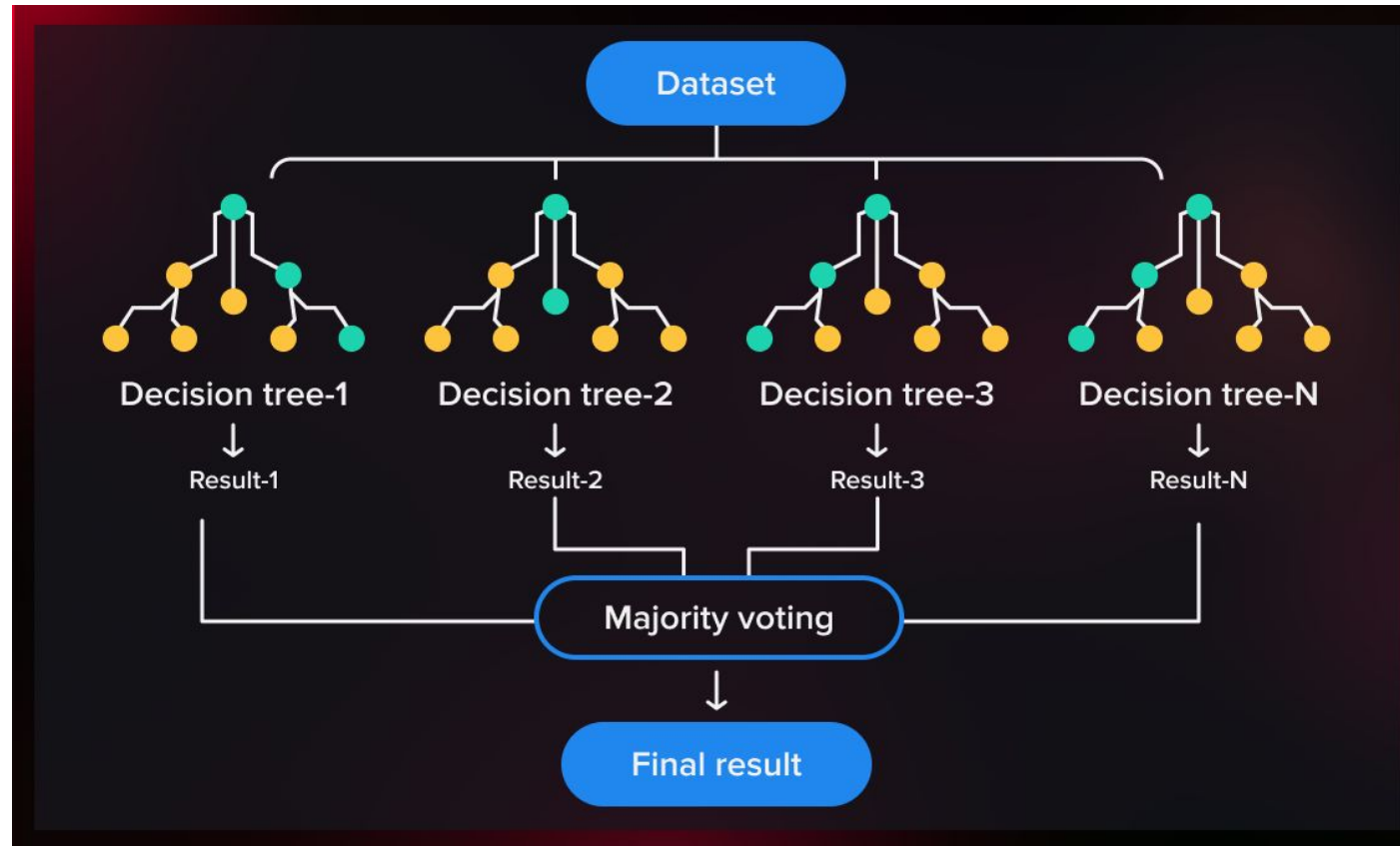


# Random Forest

---

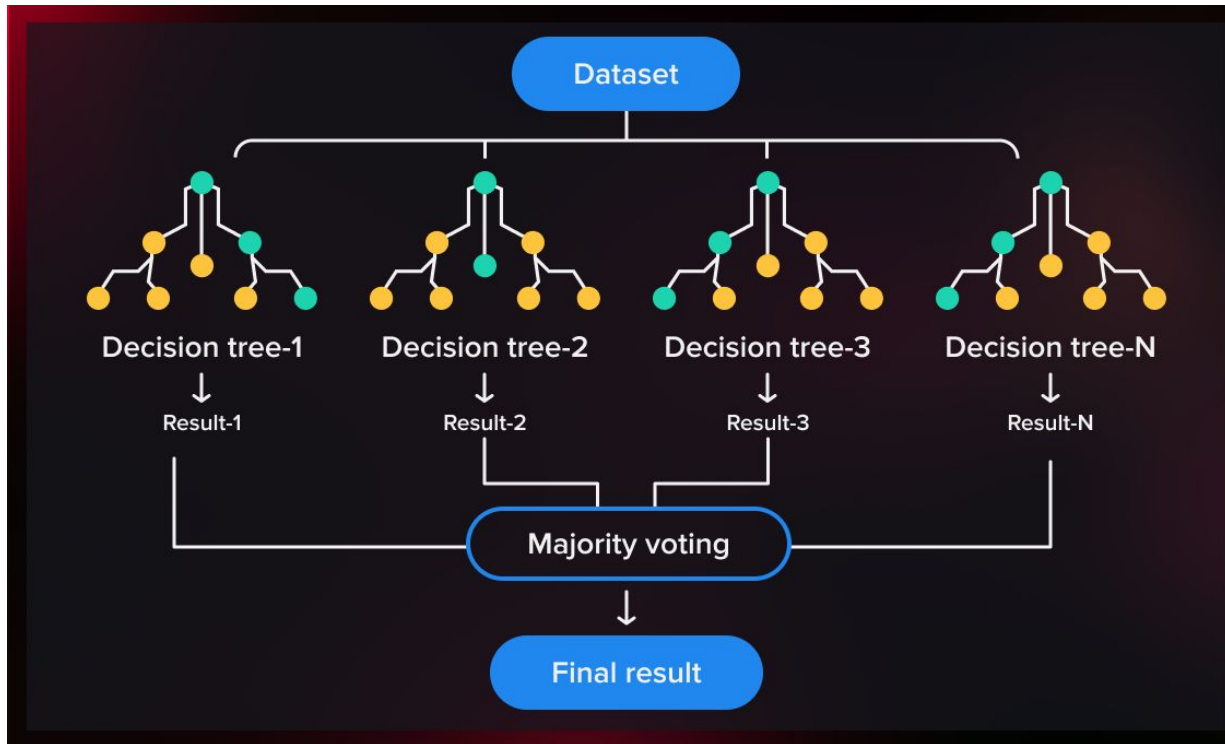
SINDHURA NADENDLA

# Random Forest

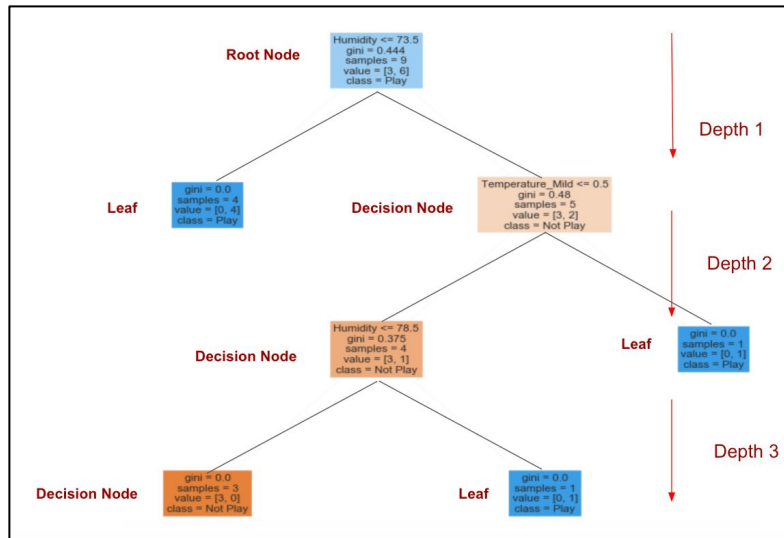


# N-estimators

How many decision trees to take ? In Below case it is 4 estimators



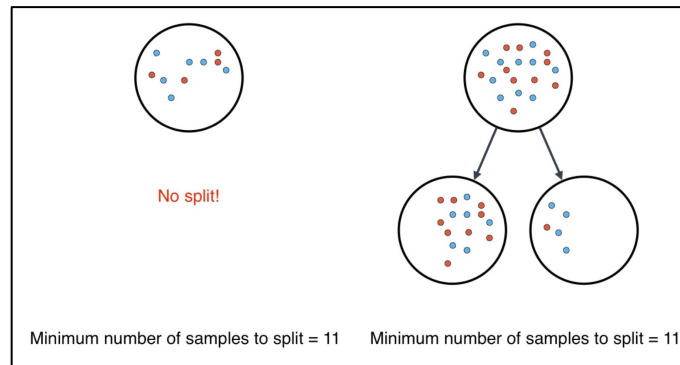
# Parameters for Random Forest



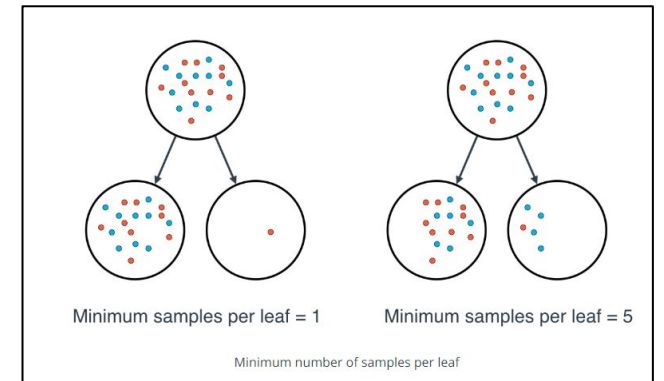
Max Depth

Regression

Criterion : squared\_error, absolute\_error



Min Samples Split



Min Samples Leaf

Classification

Criterion : gini, entropy

# Some more classification metrics

		Predicted			
		+	-		
Actual	+	TP Type II error	FN Type I error	Sensitivity (recall) TP/●	False negative rate FN/●
	-	FP Type I error	TN	False positive rate FP/●	Specificity TN/●
		Precision TP/■	False omission rate FN/■	Accuracy ( TP + TN )/( ● + ● )	
		FDR FP/■	Negative predictive value TN/■	$F_1$ score $2TP / ( 2TP + FP + FN )$	

# Scoring Metrics

---

[https://scikit-learn.org/stable/modules/model\\_evaluation.html#scoring-parameter](https://scikit-learn.org/stable/modules/model_evaluation.html#scoring-parameter)

```
In [57]: 1 import sklearn  
         2 sklearn.metrics.get_scorer_names()
```

```
Out[57]: ['accuracy',  
          'adjusted_mutual_info_score',  
          'adjusted_rand_score',  
          'average_precision',  
          'balanced_accuracy',  
          'completeness_score',  
          'explained_variance',  
          'f1',  
          'f1_macro',  
          'f1_micro',
```

# Thank you

---

PING ME ON SKYPE FOR ANY QUERIES